DOCUMENT RESUME

ED 396 084 CE 070 660

AUTHOR Mikulecky, Larry; And Others
TITLE Assessment Approaches and Impact Results in Workplace Literacy Programs.
INSTITUTION Indiana Univ., Bloomington. School of Education.
SPONS AGENCY National Center on Adult Literacy, Philadelphia, PA.
PUB DATE 95
NOTE 29p.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Adult Basic Education; *Adult Literacy; Data Collection; *Evaluation Criteria; Information Sources; *Literacy Education; Program Evaluation; *Systems Approach; *Workplace Literacy

ABSTRACT

Employers, unions, and government agencies are placing increasing emphasis on evaluation of workplace literacy programs. Because workplace literacy programs are generally short, focused on workplace-specific literacy tasks, and small, the standardized tests that have traditionally been used to assess school programs are inappropriate for evaluating them. Rigorous evaluation of workplace literacy programs is still relatively rare. A few programs have demonstrated that systematic evaluation of workplace literacy programs is possible and beneficial in improving programs and providing evidence of program effectiveness in support of requests for continued funding/expansion. The most effective approach to program evaluation is a systems approach in which formative and summative evaluation are combined and literacy task analysis is used to custom-design assessment measures. A refined model for evaluating workplace literacy programs has been developed that defines literacy broadly and that calls for assessing not only improved performance with a variety of literacy tasks but also literacy-related changes in lifestyle and in learners' self-perceptions and aspirations. A workplace literacy program's impact on job productivity can be measured in several ways, including by gathering information on individual employees and teams and by conducting job-related performance ratings. (Contains 27 references.) (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 396 084

Assessment Approaches and Impact Results in Workplace Literacy Programs

Larry Mikulecky, Paul Lloyd & Jamie Kirkley

School of Education,

Indiana University,

Bloomington, Indiana

A research project funded by the National Center on Adult Literacy

University of Pennsylvania, Philadelphia, PA

1995

CE 070660

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

P Lloyd

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

CONTENTS

Abstract	3
Introduction	4
I. The Challenges of Evaluating Workplace Literacy Programs	5
II. Current State of Workplace Literacy Evaluation in the U.S.	5
III. Effective Evaluation Methods and Practices	8
IV. A New Model of Evaluation	12
V. Assessing Program Effects on Productivity	21
Conclusion	25
References	27

ABSTRACT

Evaluation of workplace literacy programs is receiving greater emphasis from employers, unions, and government agencies. Evidence of effectiveness—and cost-effectiveness—is rapidly becoming a standard expectation for funding and refunding of any program competing for limited resources. This report describes systematic methods for evaluating workplace literacy programs and places special emphasis on multiple areas of program impact (i.e., literacy practices, processes and abilities, beliefs and plans, as well as impact on productivity), and also on multiple sources of information (e.g., interviews, questionnaires, job literacy scenarios). The middle section of this report compares learner gains on such measures with instructional practices in ten workplace literacy courses. The report concludes with suggestions of ways for measuring intellectual capital and the effect of a workplace literacy program on job productivity. These include employee job rating scales and a formula for calculating the utility of a workplace literacy program.

INTRODUCTION

For both educational and economic reasons, evaluation is growing in importance for workplace literacy programs. In the 1990s, workplace literacy programs find themselves in a new business climate. Programs are facing a more competitive market for receiving funds, and funding sources are setting more stringent guidelines for accountability. Programs are being asked to provide evidence of their own effectiveness and improvement.

Unfortunately, evaluation has not been a high priority in many workplace literacy programs. The research literature shows that only a small percentage of programs report any form of rigorous program evaluation or documentation of learner gains, impact on productivity, or detailed descriptions of program practices. This is partly because there are legitimate problems in collecting evaluation data. Factors such as short courses, open enrollment, and a lack of appropriate, standardized assessment methods make it difficult to collect accurate data on program effectiveness.

However, there are systematic methods for evaluating workplace literacy programs. By using these methods, one can assess the success of instructional methods and processes, whether partners agree about program goals, whether resources are available and used properly, and whether the instruction is being delivered in a way that follows good instructional practice and meets program expectations.

Looking to the future, it becomes obvious that the U.S. will have many people in need of education and training. We will need a population who can solve complex problems, communicate with others, manage large amounts of information, and function in a high performance workforce and society. However, need is no longer enough to justify program expenditures. Funders want to know that they are getting value for their money. Therefore, evaluation is becoming more critical for workplace literacy programs for both educational and economic reasons. Evaluation not only reveals what works now, it also reveals what improvements should be made to future educational programs.

This report provides details about methods of evaluation appropriate to workplace literacy programs. We encourage you to read further to find out more about the current and future state of workplace literacy evaluation.

I. THE CHALLENGES OF EVALUATING WORKPLACE LITERACY PROGRAMS

Evaluation of workplace literacy programs is a relatively new area, with its own unique set of problems and challenges. Program courses are usually short (30 - 40 hours), curricula often focus on workplace specific literacy tasks, and groups of learners tend to be small (10 - 15). Therefore, assessment methods traditionally used for schools—standardized tests of generalized reading abilities administered to hundreds of students—are inappropriate in this context.

Standardized reading tests measure general literacy ability and are usually used in schools to assess changes following daily instruction over a full school year. Such tests are rarely sensitive enough to assess changes over shorter time periods. Given the brevity of workplace literacy courses, their curriculum is necessarily limited and often concerns specialized workplace skills and competencies, rather than a broad range of general literacy skills. In this situation, standardized tests of general reading ability are unlikely to show much change in learners' performance, and are not often used by program providers. In order to show what learners have gained, workplace course evaluations need to use measures custom-designed to assess what has been taught. This may involve measuring the learners' ability to use workplace reading materials and to perform their jobs more effectively. An obvious measure of the latter is a worker's job productivity, but few organizations keep data of this kind on individuals—the smallest unit tends to be the work team. So the most practical measures of individual workplace competence are employee rating scales (usually filled out by supervisors) and job-related reading scenarios. And both of these need to be custom-designed for particular workplaces.

However, solving the problem of assessment in one workplace generates problems when one tries to compare gains made at different sites. The workplace-specific measures, while providing useful information on each workplace are necessarily different from each other, and this makes it difficult to compare programs across the board. Because of the various difficulties associated with evaluating workplace literacy programs, not many programs are using appropriate evaluation methods and practices. Therefore, it is difficult to compare programs and determine the most effective practices.

II. CURRENT STATE OF WORKPLACE LITERACY EVALUATION IN THE U.S.

In 1991, Kutner, Sherman, Webb & Fisher reviewed workplace literacy programs funded by the U.S. Department of Education to determine the elements of effective programs. The authors examined 29 of 37 projects funded by the National Workplace Literacy Program to determine which programs were effective and merited further examination in order to identify components of effective programs. The authors reported that:

> "Due to the absence of quantitative data necessary to identify particularly effective projects (i.e., improved productivity, low participant attrition, or improved test scores), study sites were recommended to OVAE staff. These sites were reported by project directors to have a high retention rate." (1991:26)

Even in federally funded workplace literacy programs, for which program evaluation was an expectation for receiving funding, it was not possible to find six programs which

had been rigorously evaluated for effectiveness. Selection of "effective" programs was based upon undocumented reports of retention from program directors.

Only a few workplace literacy programs described in the pre-1990 research literature reported any form of rigorous program evaluation or careful documentation of learner gains, impacts on productivity, and detailed descriptions of effective program practices. (See, for example, Sticht (1982), Mikulecky & Strange (1986), Hargroves (1989), Haigler (1990) and Philippi (1988, 1991). However, these examples are atypical.)

In another survey, Mikulecky and d'Adamo-Weinstein (1991) also observed that the majority of workplace literacy programs described in the available research literature tended to report no rigorous evaluation data. Many programs which did report evaluation data simply provided superficial information limited to surveys of learner satisfaction and anecdotal reports of effectiveness. Occasionally a pre and post administration of a standardized reading test (usually the Test of Adult Basic Education—TABE or the Adult Basic Learning Examination—ABLE) provided an indication of learner gain in general reading ability. Only a few evaluations provided follow-up data on the impact of programs on learner job performance, retention, or earning power.

In 1995, Mikulecky, Lloyd, Horwitz, Masker, and Siemantel (in press) surveyed the state of workplace literacy in the 1990s. They examined 121 workplace literacy program reports entered into the ERIC database between 1990 and 1993. They investigated the current provision of programs, both statistically and through program profiles. Few program reports said much about evaluation, and those which did rarely mentioned objective, quantifiable data showing what learners gained from attending the program.

Mikulecky et al examined program reports to see which evaluation methods were used and documented how programs used them. Four evaluation areas were examined: gathering objective, quantifiable data; measuring participant outcomes; determining the effects of instruction on jobs, and systematically evaluating the program for ongoing development.

Gathering Objective, Quantifiable Data

Mikulecky et al found that projects typically kept track of demographic data such as gender, race/ethnicity, family data, years of education, and employment data. These data are important but do not measure program effectiveness; they simply contribute to an understanding of the population reached.

Some programs attempted to gather data on program effectiveness, but much of this data was limited in its usefulness for various reasons. For example, several programs limited evaluation to end-of-program satisfaction ratings. Such programs rarely used pre- and post-measures to assess growth. Even when pre- and post-test data (e.g., test scores) was gathered, it was often reported for only some of the participants, without an adequate explanation of what became of the other scores.

There are legitimate problems in collecting this kind of data. Short programs (30 to 40 hours) are not likely to produce significant gains on tests designed to measure general literacy in grade level equivalents. Programs that are long enough (100 hours or more) to effect such changes very often have open-enrollment and open-exit policies. Learners can choose to leave before post-measures are administered, making collection of pre- and post-test data difficult.

Measuring Participant Outcomes

There are exceptions to this general trend of inadequate evaluation. Some programs systematically evaluated in several areas using multiple measures. For example, the

Training Opportunities Project in Greenville, South Carolina (Greenville Technical College, 1993) did a particularly noteworthy job of measuring the impact of training in three areas: skill level achievement, job productivity, and personal growth and development. In addition to standardized tests, anchored supervisor ratings were developed to measure dependability, team work, self initiative, job knowledge, adaptability, work pride, making decisions, leadership, and taking responsibility. Similar instruments were developed to measure changes in self-esteem and personal growth, math skills, and reading and language art skills. This type of comprehensive measurement is desirable, but not common among the programs surveyed.

Effects on Jobs

Job performance changes were most frequently measured using rating scales. Actual performance data measuring these factors were rarely reported except in projects whose primary focus was training for new employment. In these cases, productivity was measured in terms of program impact upon finding a new and appropriate job. Also, where applicable, some programs tracked whether learners met certification and licensing requirements.

Systematic Evaluation for Ongoing Program Development

There were not many models of how evaluation was used for ongoing program development. However, one good example was provided by a program designed for employees in several small businesses (custodial workers, truck drivers, daycare workers, and employees in a small manufacturing plant) in San Marcos, Texas (McBride et al, 1992). This program used a number of methods and measures to provide information for both formative and summative evaluation (these types of evaluation are further discussed in Section III). For the formative evaluation, an outside evaluator examined the program to determine if there was a shared understanding and agreement about the program's stated goals among its stakeholders. Classroom observations and examinations of resources determined the extent to which the program's instruction matched the goals of the program. Feedback received helped the providers adjust and refocus the program's goals.

For the summative evaluation, effectiveness, in terms of learner gains, was assessed using a variety of indicators before instruction, during instruction, and after instruction. Instructors interviewed workers at the outset of each course to develop an Individual Education Plan which was used both to adjust curriculum goals and to monitor learner progress. A weekly evaluation form filled out by students enabled the teacher to stay abreast of student perceptions and correct problems early in instruction or as they came up. The anonymous, written format of this weekly evaluation also provided additional writing practice for program participants. A final written evaluation form at class end and an exit interview addressed transfer from the classroom to literacy requirements at work and at home. The project managers also used more traditional pre and post measures to document reading and writing progress. They tracked retention rates and time spent on independent practice outside of class. They also initiated development of a portfolio-based qualitative assessment which program instructors felt would provide more effective and sensitive information to monitor learner gains and facilitate revision of instruction.

Although evaluations like the one described above are not typical, they do demonstrate what is possible. Many workplace literacy programs do not survive past the first few years. One reason for this is lack of information about program effectiveness and lack of information enabling program improvement. When competing for resources with other areas of a business, programs without systematic evaluation data are at a decided disadvantage.

Summary

Only a few workplace literacy programs described in the research literature report any form of rigorous program evaluation or careful documentation of learner gains, impacts on productivity, and detailed descriptions of effective program practices.

In 1995, Mikulecky et al (in press) surveyed 121 workplace literacy program reports entered into the ERIC database between 1990 and 1993, examining which evaluation methods were used and documenting how each program used them. Few program reports said much about evaluation. Projects typically kept track of demographic data, but these do not measure program effectiveness. Many programs also kept data on program effectiveness, but much of this data was limited in its usefulness for various reasons. Job performance data were rarely reported except in projects whose primary focus was training for new employment. Also, there were not many models of how evaluation was used for ongoing program development.

Systematic program evaluation is possible as has been demonstrated by a few programs. It is also important. Workplace literacy programs which last must compete for resources with other divisions of a business. Systematic evaluation data can help program providers continually improve and can provide evidence of program effectiveness to support requests for continued funding and expansion.

III. Effective Evaluation Methods and Practices

Many people think of evaluation as something one does at the end of instruction. However, the evaluation plan should be designed at the same time the program is being designed. This helps ensure a more accurate, complete, and systematic evaluation process.

When evaluating workplace literacy programs, it is most effective to use a combination of two types of evaluation -- formative and summative. The purpose of formative evaluation is to examine the instructional methods and processes while training is still under way so that changes can be made. The purpose of summative evaluation is to assess the total effectiveness of the instruction or program. (Details for conducting such evaluations, which are not given here, are available in Mikulecky, L., Lloyd, P., Kirkley, J., & Oelker, J. (1995).)

Formative Evaluation

Formative evaluation takes place during beginning and middle stages of program operation. A formative evaluation assesses the effectiveness of several areas of the workplace literacy program: whether partners agree about program goals, whether resources are available and used properly, and whether instruction is being delivered in a way that follows good instructional practice and meets program expectations. The purpose is to identify problem areas, and make changes while they are still possible and productive.

Program goals should be shared and understood by all stakeholders, including such groups as teachers, learners, managers, union officials, etc. Interviews, analysis of memos and planning documents, and early program observations often reveal that significant differences about program goals and priorities exist among funders, supervisors, instructors, materials designers and learners. Evaluation feedback during early program stages often initiates discussion and necessary clarification among program planners and participants. In some cases goals are expanded or refined, and in some cases new program providers are sought.

Formative evaluation should also address the program's resources. Resources include the expertise of key personnel, the available instructional space and materials, as well as the time available for instruction. Early examination of resources sometimes reveals that resources are insufficient to accomplish the goals espoused by program planners. In this case, planners need to regroup and decide how best to use what is available. Sometimes this means finding more resources. More often it means narrowing and redefining goals to match the resources available. Information about resources can be gathered by examining program facilities and from interviewing key program personnel.

Examining the processes used in a workplace literacy program can be accomplished through classroom observation, examining learner records and assignments, and interviewing learners and instructors. Information from such interviews can help determine if learning activities and time allocation match program goals or if learning time is insufficient to meet these goals.

We know from research that literacy improvement takes a significant amount of time and that general literacy instruction is not very effective for workplace applications (Sticht, 1982). Observation of classroom instruction, materials, and schedules sometimes reveal potential problems with the learning processes and methods offered by the program. Examples are: 1) Providing insufficient learner practice time with literacy or allocating too much class time to rapport-building discussion, 2) Teaching general reading instruction only rather than paying sufficient attention to job specific skills , 3) Using only school books, off-the-shelf materials, or sometimes materials and activities selected because the instructor has found them useful in other settings, and 4) Providing little feedback to learners about their accomplishments (sometimes instructors do not or cannot comment upon what individual learners can and cannot do).

Effective programs typically instruct learners using a goodly amount of workplace related instructional activities and real or modified workplace materials. Even when more general approaches or materials are used, effective teachers relate the instruction to learner and workplace needs. If instruction is not related to the workplace, it is because the program has simply elected to use a workplace classroom to address general literacy goals. In effective programs, no matter what the goal, sufficient learner practice time is available to allow reasonable expectation of success. Some effective programs even manage to expand literacy practice time through homework and increased practice on the job.

If problems exist with any of these areas (i.e., goals, resources, processes), formative evaluation can reveal these problems. The evaluator can recommend changes that could improve chances for success and report on areas where the program is successful.

Summative Evaluation

Summative evaluation is designed to assess how well the workplace literacy program has succeeded. The summative evaluation of the impact of workplace literacy programs is best performed using a combination of standard assessment tools and custom-designed measures. The custom-designed measures usually reflect the types of reading and writing done on the job and in training courses. In addition, they can focus on special objectives central to the workplace literacy program (e.g., increased productivity, comprehending safety information, increased participation in voluntary training).

Assessment is often accomplished through use of formal standardized tests, informally constructed tests related to the workplace, questionnaires related to literacy practices, and interviews with learners and supervisors. In addition, company records and ratings on productivity, safety, attendance, and enrollment in subsequent classes can expand the

evidence available for assessing program impact. (A more detailed discussion is available in Mikulecky and Lloyd, 1993 & 1994.)

Because the impact of instruction is measured by *changes* in learner abilities and practices, well-evaluated workplace literacy programs gather baseline data before instruction begins. Data is typically gathered on the reading abilities, practices, and beliefs of learners. In addition, pre-program data is gathered on worker productivity or any other goal espoused by the program. Data-gathering is accomplished using formal tests, informally constructed tests related to workplace expectations, questionnaires, and interviews with learners and sometimes with supervisors. The nature and extent of the instruments employed are determined by program goals: it is important that the assessment measures and company records used relate directly to those program goals.

Gathering pre-program base-line data is key to evaluation. This information establishes a base for later comparisons with end-of-program performance. At the end of the program, all learners are once again assessed using the same instruments. In some cases, it is possible to compare the performances of learners in a workplace literacy program to those of a control group of comparable employees who have not yet been able to receive workplace literacy training. To do this, the control group takes pre and post assessments which parallel the instructional group.

It is important to determine exactly what will be evaluated early in the planning process. As curriculum is being designed, assessments to use as pre and post measures should also be developed. These assessments should reflect as closely as possible the objectives of the curriculum being taught, and be relevant to both teaching and jobs. For example, if one program goal is to improve the accuracy of filling out work order forms, then an appropriate assessment would involve a job scenario based on one of these forms. (See Section IV for more details about such scenarios.) It is important to test what is being taught, or there will not be accurate data on learner gains. So that results can be directly compared later, administer the same assessment for both the pre- and post-test. When the same assessments are used, it is important to allow sufficient time for learners' memories of the details to decay (i.e., 6 weeks or more) and for teachers not to provide feedback on correct answers until after the post-test.

Standardized tests such as the Test of Applied Literacy Skills (TALS) and the CASAS Life Skills Assessment can sometimes be used as part of summative evaluation. Those measures are most effective when curriculum matches the types of life skills materials and tasks used on these measures. Sometimes only a single sub test is needed (e.g., document or quantitative). As with the custom-designed measures discussed above, standardized tests should be administered as pre- and post-assessments so that results can be compared directly. The TALS and CASAS are often appropriate for workplace education because each uses realistic tasks (e.g., adding up a bank deposit slip, reading a table).

Program goals determine the types of information gathered to assess program impact. For example, if a program goal is to improve the ability of learners to perform more effectively in quality assurance groups, evidence needs to be gathered on such performance before and after instruction. If instruction is supposed to have a positive impact on learner reading habits at home and at work, these, too, need to be assessed before and after the program.

Typical goals for workplace literacy programs include improved learner literacy abilities, improved literacy practices at work and elsewhere, changed learner beliefs about literacy, self, and education, and improved learner productivity on the job. Many programs have goals that go beyond increased literacy skills to achieving changes in literacy practice

outside of class, expanding employee educational aspirations, and increasing employees' senses of personal effectiveness with literacy.

In addition to measuring gains in literacy ability and life style, summative evaluation also examines the impact the program had on productivity in the workplace. Productivity indicators may show improvement among those who attended training. Some types of productivity indicators are supervisor ratings, company records, and job scenarios. In this and other areas, it is important to use multiple measures rather than relying on one type of measure. For example, to measure customer service representatives' skills in dealing with customer complaints, learner gains could be assessed using a pre/post job scenario of a typical complaint, pre/post supervisor ratings of workers' ability to deal with complaints, and company records of complaint handling for a period of time before and after class attendance.

Finally, summative evaluation assesses how well the program fulfilled its stated goals in terms of providing satisfactory service, producing new reusable curriculum materials, and fostering continued learning among employees. Documentation in this area can include a list of goal statements and related program achievements, files of lesson plans and materials, and statistics on later courses attended by employees.

When conducting either a formative or summative evaluation, it is important to gather information from a variety of sources. Use information gained from interviews, observations, and any appropriate records. This will help ensure a more appropriate and complete assessment.

Using Literacy Task Analysis to Custom Design Assessment Measures

Literacy task analysis is a way of identifying those aspects of job tasks which require reading and problem solving. These analyses are performed using a combination of observations of workers, interviews with top performers, and gathering samples of print used in the workplace and training classes. The goal is to determine the mental processes used by top performers as they solve problems and complete tasks which involve literacy. This information can be used to construct both test scenarios and instructional materials. It is important that these two be developed together, so that tests can assess what learners are really taught and both can be linked directly to the workplace.

Observations and interviews with supervisors and workers are used to identify the areas in which performance needs to be improved. Prime targets for literacy task analyses are tasks where basic skills problems cost money or threaten health and safety. Other tasks can be identified by noting changes in the workplace (e.g., new technology, changed jobs or promotions) which confront some workers with new and sometimes troublesome literacy tasks.

A good deal has been written about how to perform literacy task analyses (see Mikulecky, 1985; U.S. Departments of Education and Labor, 1988; Drew & Mikulecky, 1988; Philippi, 1988 & 1991, Norback, Rosenfeld & Wilson, 1994, Norback, in press). Most techniques involve determining the elements of a task and the strategies (both visible and mental) employed to accomplish the task. For example, filling in forms in some quality assurance procedures involves the elements of reading two-column charts, computing using decimals, knowing special vocabulary and abbreviations, and being able to summarize sequences of events. Within each of these elements, top performers employ a variety of strategies—skimming, estimating, and interpolating.

Materials and information gathered during literacy task analyses can be used to develop instructional materials as well as to develop custom-designed assessment instruments for

workplace literacy programs. Examples of such instruments (i.e., literacy scenarios) are discussed in Section IV.

In addition to their use in a pre-test to establish base-line data for assessment, job scenarios can be used at the beginning of a program to diagnose areas of learner difficulty. If the information on the scenarios is also part of a curriculum, the scenarios can provide instructors with valuable information. For example, if a learner consistently has difficulty with inference questions across scenarios, the instructor can adjust instruction to provide more guidance and practice in this area. The instructor should not, however, provide detailed feedback to learners about their performance on particular scenarios if the program intends to use those scenarios again as a post-test to assess learner gain and program effectiveness.

A test can be used a second time to indicate learner growth if the learner has not been taught or given feedback using the actual test. It is also important that sufficient time has passed between pre- and post-tests (six weeks is usually sufficient) for detailed memory to decay. If such time is not available, it is possible to develop two very similar tests and establish the comparability of the two scenarios by noting how a pilot group scores on them. This is a fairly lengthy procedure, but worthwhile if the tests will be used with many learners for several years. Once comparability has been established, the two forms of the scenario can be used as pre- and post-measures. However, using the same scenarios for both tests provides a more reliable means of establishing comparability.

Summary

Evaluation is an integral part of any program because it shows the effectiveness of the instruction and the program. The evaluation plan should be designed at the same time the program is being designed to ensure a more accurate evaluation process.

When evaluating workplace literacy programs, it is most effective to use a combination of formative and summative evaluation. Formative evaluation assesses the effectiveness of several areas of the workplace literacy program: whether partners agree about program goals, whether resources are available and used properly, and whether the instruction is being delivered in a way that follows good instructional practice and meets program expectations. Problem areas can be identified and changes can be made early in the program.

Summative evaluation assesses the total effectiveness of the instruction or program and is best performed using a combination of standardized and custom-designed measures. It is important to remember to use multiple measures of testing, test what you teach, and use pre and post assessments to determine learner gains. Using a combination of custom-designed measures along with standardized tests will provide a useful profile of learner gains. Examples of customized measures include job scenarios and performance ratings. Custom-designed measures can be developed by performing a literacy task analysis to identify job tasks which require reading and problem solving. It is important to establish base line data by giving pre-tests. Those pre-tests can later be compared to post-tests to evaluate learner gain and program success.

IV. A New Model of Evaluation

In 1995, Mikulecky and Lloyd (in press) studied ten groups of learners in workplace literacy programs at several different sites. This three-year study was designed to determine the feasibility of developing a workplace literacy assessment model which could produce

information useful for local program providers and for funders and researchers interested in aggregating data across a group of programs. They gathered data on the impact of programs in the area of learner gains and workplace improvements and then developed and refined a model for evaluating workplace literacy programs.

Assessing a Broader Conception of Adult Literacy

The conceptual framework for this workplace literacy impact assessment model was based upon a broad understanding of literacy. In addition to assessing improved performance with a variety of literacy tasks, the model also assessed changes in lifestyle related to literacy and changes in learners' self-perceptions and aspirations in relation to literacy. The model owes much to ideas presented and developed by Lytle (1990) in long-term case-studies of changes in adults experiencing success in adult literacy programs. It was also influenced by ideas about perceptions of self-efficacy developed by Bandura (1986) and others who have studied why some individuals outperform others who have similar tested abilities.

Lytle (1990) has suggested that performance measures (tests and exercises) miss a good deal of important information about adult literacy learning. She has examined the literacy growth of adults enrolled in adult literacy programs and used extensive observations, interviews, and learner journals to track changes which took place as learners spent a year in adult literacy instruction. Lytle (1990) found that in addition to gains in literacy skills, adults tend to make changes in what they believe, how they behave, and in their aspirations. These changes are intertwined and seem to influence each other. For example, new understandings about the process of how reading and writing work may influence beliefs about what the learner can accomplish. Changes in aspirations resulting from an increased sense of effectiveness sometimes lead to increased literacy practice and increased competence. Lytle suggests several dimensions which constitute a fuller understanding of adult literacy and adult literacy growth. These dimensions are learner *beliefs* about literacy and themselves, learner literacy *practices*, the literacy *processes* employed by learners while reading, and the *plans* a learner has which may involve literacy use. Programs and assessments which focus merely upon performance are likely to miss the complex, intertwined mixture of changes in:

- self-perception,
- literacy life-style,
- understanding of the literacy process, and
- hopes and aspirations,

which are necessary elements in improving and sustaining literacy performance.

Bandura (1986) has written extensively about the influence upon performance of self-perception of one's effectiveness. His concept of "perceived self-efficacy" has been used to examine performance of children and adolescents in school, as well as the performance of adults in a variety of life situations. Adult literacy researchers (Van Tilburg & DuBois, 1989; Bean, Partanen, Wright, & Aaronson, 1989) have noted distress leading to poor performance and abandoning of programs by adults receiving literacy instruction. Perceived self-efficacy (based upon accurate feedback) is particularly important in relation to adult literacy learning. While learners with low senses of personal effectiveness tend to subvert their own efforts with self-doubt and excuses for quitting, learners with higher senses of effectiveness often perform successfully and continue to learn as a result of persistence.

In the workplace literacy program impact model developed for the National Center on Adult Literacy, Lytle's conception of learners' beliefs about literacy is expanded to focus more specifically upon beliefs about personal effectiveness with literacy. Her conceptual framework has been adapted to test the importance of these aspects of adult learning:

beliefs, practices, process and plans, and in order to seek out ways to enhance learning. Information about these dimensions of learner literacy were gathered in tests of the model (Mikulecky & Lloyd, 1993) using a combination of questionnaire items, interview questions, and requests that learners explain their literacy strategies or processes while simulating job tasks.

Beliefs

In the interview, learners were asked to describe themselves as readers and writers and to describe someone who seemed to be very good at reading and writing. They were also asked to provide reasons for their answers. Changes in these beliefs are likely to precede changes in literacy abilities. Sample questions from the interview follow below.

Beliefs

1. Describe someone you know who is good at reading and writing. What makes you choose this person?

2. How good do you consider yourself to be at reading and writing? What makes you think so?

3. Describe how you would like to be in terms of reading and writing. (*Probe* : Can you give me some examples?)

Practices

Learners were asked orally and in the questionnaire for information about the types of reading and writing they do on the job and off the job . They were asked to rate the difficulty they had in reading each item on a list that included books, signs, training manuals, pay stubs, charts and cartoons. They were also questioned about the frequency of their literacy-related activities: how often, for example, they read a newspaper, made a shopping list or visited a library, as well as how many books they owned. Information was also sought about literacy practices in workplace situations which ranged from departmental meetings to handling broken equipment, from reading instruction manuals to reading a health insurance policy. (Sample interview and questionnaire items follow below.)

Practices

Interview item

Tell me the sorts of things you read and write away from work during a normal week. (*For probe, ask:* "Can you give me more examples?")

Questionnaire items

1. First check only the things you've read in the past month.

   Now go back and rate your ability to read the items you've checked.

	poor				excellent
__ local newspapers	1	2	3	4	5
__ classified ads	1	2	3	4	5
__ telephone bills	1	2	3	4	5
__ TV guide listings	1	2	3	4	5
__ magazines	1	2	3	4	5

2. In the last 7 days how many times have you read a newspaper?

0 1 2 3 4 5 6 7 8 9 10+

3. You talk a lot in team or department meetings, asking questions or sharing ideas.

very like me 1 2 3 4 5 very unlike me

Process and Ability: Job Scenarios

Simulations and scenarios can be constructed by using actual materials from the workplace to assess the job literacy abilities and literacy processes of workers. Information and materials gathered during a literacy task analysis form the basis for constructing job-like scenarios in which the learner reads and makes decisions based on written materials. Scenarios are usually constructed to reflect a range of material types (i.e., prose, documents, graphic material), and sometimes involve both reading and computation. If the range of learner reading abilities is likely to be wide, it is useful to construct scenario questions which range from fairly easy to fairly complex, so that all test-takers can experience success at some level.

For full range testing purposes, it is recommended that scenarios include process questions, factual questions, inference questions, and application questions. Process questions determine how the reader reads a passage: that is, the range and sophistication of reading strategies employed (e.g., skimming, focusing, asking questions) and whether the choice and use of reading strategies improved as a result of instruction. Factual questions should have answers based directly on the reading material, answers to inference questions can involve deductions from several places in the reading, and application questions should

relate the reading to the interviewee's background knowledge. (See the following example.)

Process and Ability
Process question

I am going to show you a newspaper article about your industry.
Explain to me how you would read this story in order to find out what the writer thinks.
Describe what you would look at. What would you be thinking about?
How would you go about reading this story? What would you do first, then next, then next?

Factual question
How many employees does ASMO have in Statesville?
(Answer: 400. *Listed in article*)

Inference question
From the information provided about products, what do all four companies have in common?
(Answer: All of them make some sort of motor. *Requires the interviewee to search for commonalities not readily apparent.*)

Application question
What company makes products closest to your job at this facility?
Why do you say so?
(Answer: Relate a product on the list to what the employee makes. *Requires the employee to sort through the information and then to apply it to his/her background knowledge.*)

Plans

Some questions in the interview sought information about the learners' plans, especially in relation to education and goals requiring increased literacy abilities. These questions asked for information about learner plans for one year, five years, and ten years ahead. Sample questions follow below.

Plans

Now I'd like to ask you about your plans. Explain how you see reading and education as part of these plans:

1. What are your plans for the next year?

2. What are your plans for the next 5 years?

3. What are your plans for the next 10 years?

How Information Was Gathered to Test the Evaluation Model

The purposes of the evaluation model study were: 1) to gather data on the instructional effectiveness in a variety of areas of a number of workplace literacy programs, and 2) to compare those results with the nature of the instructional practices in the programs. This comparison was intended to reveal specific connections between teaching and learning practices and learner gains in the several areas related to workplace literacy in the National Center for Adult Literacy (NCAL) workplace literacy impact model . The model focuses its assessment on changes in learners' beliefs about personal effectiveness with literacy, changes in learners' literacy practices, learners' literacy improvement with general and workplace materials, and changes in learners' goals.

Project personnel worked with onsite coordinators at each company to develop a mix of shared and customized instruments using model guidelines. These instruments were designed to assess the effect of the literacy courses on learners' own literacy behaviors and on their work competency. For the first of these, Lytle's literacy model "Beliefs, Practices, Process, Plans" was used as a basis for structuring measurement of learner change (Lytle, 1990). During structured pre and post interviews, learners were asked about themselves and their abilities in relation to literacy, about their reading and writing practices, about how they read print materials, and about their future educational plans. Pre-test data were gathered at the start of each course and post-test data toward the end. All data were then sent in to the project for analysis. In addition, discussions with instructors, classroom observations, and analysis of curricular materials provided a foundation for rating instructional emphases at each worksite.

In order to analyze the results of the ten groups of learners, rating schemes were constructed to summarize the characteristics of each course and its curriculum, in such areas as emphasis on workplace examples, reading and writing intensity, and discussion of literacy beliefs and plans. Comparisons between these course characteristic ratings and learner gain scores have been used to point up program characteristics which produce learner gains. The areas of learner gain were:

1. Practices at work
2. Practices away from work
3. Reading process
4. Scenario performance
5. Beliefs (literacy self-efficacy)
6. Plans

Following are two examples of factors used for rating course characteristics.

Example 1:

Workplace Orientation of Curriculum and Materials:

0 have no direct connection with the workplace
1 use workplace examples occasionally
2 use workplace examples sometimes (20-30% of time)
3 use workplace examples much of the time (50-60% of time)
4 are connected mainly to the workplace (70-80% of time)
5 are connected entirely to the workplace (90-100% of time)

Example 2:

Discussion of literacy beliefs and plans:

0 does not occur in this course
1 occurs occasionally/incidentally in this course
2 occurs as a deliberate part of this course
3 occurs moderately often in this course (every other session)

4 occurs often in this course (most sessions)
5 occurs very often in this course (every session)

Results of Using the Workplace Literacy Impact Model Across Several Sites

The technique of analysis of variance (ANOVA) was used to compare the gains of those groups having high ratings on particular course characteristics (such as workplace orientation or discussion of beliefs and plans) to those having low ratings on those characteristics. This indicated which course characteristics play important roles in determining whether learners make gains in each area and which program practices lead to success and in what areas.

In addition, patterns across courses can suggest a threshold level of the activity related to that characteristic (i.e.,use of workplace materials or discussion of literacy processes) required to produce measurable learner gains in various areas of the workplace literacy program impact model. Though this process is only a beginning attempt to identify patterns across programs, it can allow some tentative insights about a desirable mixture of course activities which will produce learner gains in a variety areas related to workplace literacy competence.

The divisions for all course characteristics are shown below in Table 1. Technical details on how high/low cut-offs were determined are available in the project technical report (Mikulecky and Lloyd, 1995).

Table 1: High and Low Ratings of Course Characteristics

Course characteristic	Division	Description
Instructional time in hours	low	up to 50 hours
	high	over 50 hours
Workplace orientation	low	little connection with workplace
	high	use workplace examples at least 20-30% of time
Discussion of literacy beliefs and plans	low	at most occasional/incidental
	high	deliberate part of course
Discussion of reading and writing processes	low	at most occasional/incidental
	high	deliberate part of course
Reading/writing intensity	low	up to 70% of time
	high	over 70% of time

The areas of learner gains (practices at work, practices away from work, reading process, scenario performance, beliefs, plans) were compared with the course characteristic ratings, and cut-points in the rating scales were identified at which the amount of learner gain changed significantly. Using these divisions into high and low ratings on the course characteristics, we have listed the following statistically significant ANOVA results.

Gains in Reading Practices and Processes

The learners with 200 instructional hours (one class at one workplace) made significant gains in reading practices away from work, but all the other learners, with 50 hours or less, did not.

The learners who spent over 70% of their instructional time reading and writing had a mean gain in self-reported sophistication in reading processes, which was over three times that of the other learners, although both the high- and low-rated learners made significant gains in this area.

Gains in Reading Ability

The learners who used workplace examples in class at least 20-30% of the time had a mean gain on the job reading scenarios which was nearly twice that of the other learners, although both the high- and low-rated learners made significant gains.

The learners who had discussion of reading and writing processes as a deliberate part of their instruction made significant gains on the job reading scenarios, but the other learners did not.

The learners who had discussion of literacy beliefs and plans as a deliberate part of their instruction had a mean gain on the reading scenarios which was nearly three times that of the other learners, although both the high- and low-rated learners made significant gains.

Changes in Literacy Beliefs and Plans

The learners who had discussion of literacy beliefs and plans as a deliberate part of their instruction made significant gains in the area of beliefs and perceived self-efficacy in relation to literacy, but the other learners did not.

The learners who used workplace examples at least 20-30% of the time had a mean gain in the area of improved educational plans, but all the other learners did not.

The learners who had discussion of literacy beliefs and plans as a deliberate part of their instruction made significant gains in the area of improved educational plans, but the other learners did not.

Conclusions

The study demonstrated the feasibility of implementing the assessment model at a variety of sites. The inclusion in the model of a mixture of common assessment instruments used at all programs and a framework for developing customized instruments for each program has allowed for the variations in job materials used and skills being taught at the different workplaces. Also, employing a broad conception of literacy makes it possible to be sensitive to program variations and to demonstrate program gains even when evaluation constraints are very limiting.

Study results also suggest the following structure for a workplace literacy course to be a success in a wide variety of areas. It should include a large proportion of time when learners practice reading and writing (70 - 80% of course time) and a substantial proportion of workplace examples (about 30% of course time). Integrated into this, but without detracting from the reading and writing practice time, there should also be planned regular discussion both of learner beliefs and plans concerning literacy and of reading and writing processes. With such a mix, the results above suggest that learners ought to make gains in

BEST COPY AVAILABLE

their reading abilities and sophistication of strategy knowledge, in their beliefs in their own literacy effectiveness, and in their abilities to plan for a future connected to literacy and education. In addition, for longer-running courses (i.e., 200 hours), changes in learners' everyday literacy practices may also be expected.

The brief duration of most workplace literacy programs and the limited degree of instructional transfer make it mandatory that program providers have clear goals for what they want to achieve in the limited time that learners are in class. Amongst the goals should be helping learners develop, through discussion and feedback, clearer senses of their own improving literacy abilities, broader senses of the literacy strategies available to them, and stronger links between what they are currently learning and future occupational and educational choices. Since time is so short in courses, instructors should also be seeking ways to extend this time beyond the classroom. One way of doing this is to use on-the-job materials in class so that learners are more likely to continue practicing outside class time. Also, encouraging learner motivation and independence is likely to lead to learners engaging more often in literacy-related activities.

Workplace literacy program impact is best measured using a mixture of standard assessment tests and custom-designed instruments. Standardized tests provide useful information about *general* reading ability, but may be misleading with regard to workplace literacy skills.

Custom designing starts with a literacy task analysis to identify aspects of job tasks which require reading and problem-solving, and in which performance needs to improve. Job scenarios can test their skills in using what they read, through process, factual, inference and application questions.

A broader conception of adult literacy learning can be assessed by seeking information about the learners' literacy beliefs, practices, processes and plans, using interviews and questionnaires.

Summary

In 1995, Mikulecky and Lloyd (in press) studied ten groups of learners in workplace literacy programs. The purposes of this three-year study were : 1) to gather data on the effectiveness in a variety of areas of a number of workplace literacy programs, and 2) to compare those results with the nature of the instructional practices in the programs. This comparison was intended to reveal specific connections between teaching and learning practices and learner gains in the several areas related to workplace literacy in the National Center for Adult Literacy workplace literacy impact model . The model focused its assessment on changes in learners' beliefs about personal effectiveness with literacy, changes in learners' literacy practices, learners' literacy improvement with general and workplace materials, and changes in learners' goals.

An analysis of variance was used to indicate workplace literacy program practices which lead to success within the various components of the workplace literacy impact model. Data was gathered by comparing the differences between the gains of those groups of learners having high ratings on particular course characteristics with those having low ratings on that characteristic. This indicated which course characteristics play important roles in determining whether learners make gains in each area. Though this process is only a beginning attempt to identify patterns across programs, it can allow some tentative insights about a desirable mixture of course activities which will produce learner gains in a variety areas related to workplace literacy competence.

The following structure for a successful workplace literacy is recommended. It should include a large proportion of time when learners practice reading and writing (70 - 80% of course time) and a substantial proportion of workplace examples (about 30% of course time). Integrated into this there should also be planned, regular discussion both about learner beliefs and plans concerning literacy and of reading and writing processes. With such a mix, the results above suggest that learners ought to make gains in their reading abilities and the sophistication of their reading strategies, in their beliefs in their own literacy effectiveness, and in their abilities to plan for a future connected to literacy and education.

V. Assessing Program Effects on Productivity

As companies become more knowledge- and information-based, they are looking for new ways to identify knowledge assets, measure their value, and enhance them through instructional programs or other means. However, knowledge and information are not often reported on balance sheets. The old ways of measuring resources and productivity no longer accurately reflect today's businesses. Today's accounting methods do not measure intellectual assets, even though their worth often exceeds that of raw materials and capital. "What's at stake is nothing less than learning to operate and evaluate a business when knowledge is its chief resource and result" (Stewart, 1994).

Companies today are investigating ways of measuring and accounting for knowledge, skills, and information. Skandia Assurance and Financial Services came up with a system for quantifying various types of knowledge, such as investments in information technology and business innovations. Along with measuring the growth of its network and the size of its accounts, Skandia also considered the interrelationships among customers, processes, and renewals. Skandia examined their process of "transforming human capital into structural capital." In their "Annual Report on Intellectual Capital in 1993," they used ratios to indicate how effectively they leverage their intellectual assets. (Stewart, 1994).

In 1993, Dow Chemical hired a director of intellectual asset management to manage 29,000 in force patents. Dow was taking advantage of less than half its patents, and this was costing the company millions of dollars in potential earnings. The intellectual asset manager came up with an active process for evaluating the potential worth of the patents (or knowledge) and tracking them. He helped Dow save over $1 million.

Along with the new ways of measuring knowledge, skill, and information come new ways to evaluate the usefulness and value of education. Companies now want to make sure they are getting their money's worth from education programs, and they are trying to find ways to measure the value of education.

"Very little research exists about the relationship of literacy to job performance. Much of what exists is sketchy and based on information obtained from studies conducted in the military" (U.S. Departments of Education and Labor, 1988: 37). The reasons for this are two-fold. First, Collino, Aderman & Askov (1988) found that, even when companies do conduct assessments of their literacy programs, the results are not usually made public. Furthermore, such assessments rarely involve a study of how productivity might be affected. Second, until recently, organizations have regarded workplace literacy programs more as philanthropic than as business enterprises. So some have not considered it appropriate to subject them to their usual cost-benefit analyses. In the 1990s, this is changing.

There are several ways to measure a workplace literacy program's impact on job productivity. One method is to examine company records, such as scrap rates, use of suggestion boxes, attendance at work or in classes, number of customer complaints, and participation in quality assurance groups. By gathering information on individual employees or on teams taking training, one can assess the impact of the program on productivity and other organizational goals.

Another measurement is job-related performance ratings. These ratings are a series of anchored rating scales that are custom-designed for a particular workplace. This allows one to determine how successfully employees can apply new knowledge on the job. Examples of job performances include reading measurements and setting up equipment, following safety rules, communicating with team members, and completing paperwork. First, determine the key areas of performances areas, and then ask supervisors (or others who are aware of individual's performance) to rate the employees both before and after the program. Both company records and employee performance ratings are examined before and after instruction to measure the program's impact on job productivity.

Following is an example of two job performance rating scales.

Employee Assessment - Overall Rating

Please rate each employee on a scale of 1 - 10 for each aspect below.

- An average employee would be rated 5.
- A top employee would be rated 8 or higher.
- A bottom employee would be rated 2 or lower.

EMPLOYEE _____
DATE _____
RATER _____

PROBLEM-SOLVING

Bottom	Average	Top
calls supervisor on minor details or continues to work when equipment is faulty	makes minor adjustments, offers solutions to problems and calls supervisor only when necessary	can analyze job situations, make suggestions and solutions which implement change

1 2 3 4 5 6 7 8 9 10

MACHINE SETTING

Bottom	Average	Top
unable to set machines correctly	usually sets machines correctly, but doesn't always check settings	sets machines correctly and checks settings thoroughly

1 2 3 4 5 6 7 8 9 10

Formula for Determining the "Utility" of Training

The renewed emphasis on measuring the worth of training has brought about several models and formulas to help calculate it. One such model is the expense model, which allows one to measure the economic consequences of an employee's behavior on an organization. This model accounts for performance differences between employees, and improved performances resulting from instruction.

Calculating the "Utility" or cost effectiveness of a training program requires:

1. An overall measure of the job performance of each employee trained and of a comparable group of untrained workers. (This could be either a supervisor rating or be based on production outcomes.)

2. A measure of the dollar value to the company of the difference between outstanding and average employees. For example, is a top employee worth 1 1/2 average employees?, two average employees? (This estimate of the standard deviation of performance is known as the "value".)

3. The expected duration of the training's effect. (Is training good for 6 months, a year, five years?)

4. The cost of the training (Sheppeck and Cohen, 1985).

An example of an overall measure of job performance could be a supervisor rating (for example on a scale of 10 or 100) or it could be determined by production outcomes. The difference between the average scores of the trained and untrained employees is standardized—divided by the standard deviation of the untrained group—to give the "performance difference".

Also needed is a measure of the dollar value to the company of the difference between outstanding and average (or average and marginal) employees; this estimate of the standard deviation of performance is known as the "value". Together with the expected duration of the training's effect and the cost of the training program, these can be used in the formula:

"Utility" =

Years duration of effect x Number trained x Performance difference x Value — Number trained x Cost per trainee

Example:

20 employees who have completed a training program are rated on average at 65 out of 100 by their supervisors compared with all other (untrained) employees, who average 50 with a standard deviation of 10. Then the trained workers are at a level of 1.5 standardized units above the untrained: (65 - 50) / 10 = 1.5, which is the performance difference.

If, in addition, supervisors estimate that an average employee is worth $18,000 to the company and an outstanding one $26,000, then an estimate of the "value" or standard deviation of employee performance is $8,000 (the difference between these two amounts). Another way of looking at this is that a top employee is worth nearly 1 1/2 average employees (i.e., $18,000 plus $8,000).

Then suppose that the cost of training is $2,000 per employee and the effect of the training is likely to last 3 years. Then, using the formula,

Utility = Years duration of effect x Number trained x Performance difference x Value — Number trained x Cost per trainee,

the utility of this program is:

3 x 20 x 1.5 x $8000 — 20 x $2000
= $720,000 — $40,000
= $680,000.

In other words, for an outlay on training of $40,000, the company's return is $720,000 in the form of more productive employees.

In some situations, it may prove difficult to assess the "value" of employees to the company. A useful approximation that can be used in such cases is that "value" is typically between 40 and 70 percent of the average yearly salary for the job in question (Sheppeck & Cohen, 1985). For example, for second level managers making an average of $30,000 a year, the value would probably lie between $12,000 and $21,000. When estimates of the "value" range widely like this, the calculation of the utility of training should be carried out with both the lower and upper estimates to obtain an indication of the program's utility. Similarly, if other parameters such as the duration of effect could vary, an approximation to the utility can still be calculated using several different values.

With companies looking more closely at their bottom lines, it is important to be able to determine not only the cost of training but also its benefits. It is also important to be able to measure it in quantifiable terms.

Summary

As companies become more knowledge- and information-based, they are looking for new ways to identify knowledge assets and measure them. The old ways of measuring resources and productivity no longer accurately reflect today's businesses. Today's accounting methods do not measure intellectual assets, even though their worth often exceeds that of raw materials and capital.

Along with the new ways of measuring knowledge, skill, and information come new ways of evaluating the usefulness and value of education. Companies now want to make sure they are getting their money's worth from education programs, and they are trying to find ways to measure the value of education. However, very little research exists about the relationship of literacy to job performance.

There are several ways to measure a workplace literacy program's impact on job productivity. By gathering information on individual employees and on teams, one can assess the impact of the program on productivity and other organizational goals. Measurements such as job-related performance ratings allow one to determine how successfully employees can apply new knowledge on the job.

With companies looking more closely at their bottom lines, it is important to be able to determine the benefits of training in quantifiable terms. The renewed emphasis on measuring the worth of training has brought about several models and formulas to help calculate it. By using various factors such as the economic consequences of an employee's behavior on an organization, performance differences between employees, and improved

performances resulting from training, one can determine, in dollars, the utility of a training program.

CONCLUSION

Ten years ago, there were relatively few workplace literacy programs. Of those few, many were considered a form of "corporate welfare" or the result of altruistic impulses on the part of visionary leaders. In the 1990s, workplace literacy programs are more plentiful and find themselves in a new climate. Business is focused upon productivity, quality, and effectiveness. Assessment and quality monitoring have become part of the expectations of all departments, and workplace literacy programs find themselves competing for tight resources. They are being asked to provide evidence of effectiveness and improvement, just like every other department.

Unfortunately, evaluation has not been a high priority in many workplace literacy programs. Research literature shows that only a handful of literacy programs report any form of rigorous program evaluation or documentation of learner gains, impact on productivity, and detailed descriptions of program practices. There are legitimate problems in collecting evaluation data. Factors such as short courses open enrollment, and a lack of appropriate, standardized assessment methods make it difficult to collect accurate pre- and post-test data. Therefore, it is difficult to systematically measure groups and compare results.

Two types of evaluation provide a systematic way of evaluating program development and outcome. These are formative and summative evaluation. The purpose of formative evaluation is to examine the instructional methods and processes while instruction is still under way so that changes can be made. Using formative evaluation, one can assess the effectiveness of several areas of the workplace literacy program: whether program goals were met, whether resources were available and used properly, and sound practices were being employed.

Using summative evaluation, one can assess the total effectiveness of the instruction or program and how well the workplace literacy program has succeeded. A summative evaluation of the impact of workplace literacy programs is best performed using a combination of standard assessment tools and custom-designed measures. Assessment is often accomplished by using a variety of measures, such as formal standardized tests, informally constructed tests related to the workplace, questionnaires related to literacy practices, interviews with learners and supervisors, company records and ratings on productivity, safety, attendance, and enrollment in subsequent classes

Well-evaluated workplace literacy programs gather baseline data on the reading abilities, practices, and beliefs of learners. In addition, pre- and post-data is gathered on worker productivity or any other goal espoused by the program. Assessment measures should relate directly to program goals. Using a literacy task analysis, custom designed assessments can be developed while the curriculum is being designed. The assessments should reflect as closely as possible the objectives of the curriculum being taught, and be relevant to both teaching and jobs. It is important to test what is being taught, or there will not be accurate data on learner gains.

The goals of workplace literacy programs are diverse and change form from worksite to worksite. The time available for instruction is usually quite brief—classes usually meet 50 hours or less. For these reasons, it is important for program planners to set reasonable

goals and expectations or what can be accomplished in the brief time available. Both curriculum and assessment should be targeted to address the high priority goals.

Today's accounting methods do not usually measure intellectual assets, even though the worth of intelligence factors in a business often exceeds that of raw materials and capital. As companies become more knowledge- and information-based, they are seeking new ways to identify knowledge assets, measure their value, and enhance them through training or other means.

Along with new ways of measuring knowledge, there are new ways to measure and evaluate the usefulness and value of education. The renewed emphasis on assessing the worth of training has brought about several models and formulas that help calculate it. The expense model allows one to determine the economic consequences of an employee's behavior on an organization. This model accounts for performance differences in employees and improved performances resulting from training. From these, one can determine the utility of training in dollars.

With more companies examining their bottom lines, it is important to be able to determine not only the cost of training but also its benefits. It is also important to be able to measure training in quantifiable terms and show how it affects job performance and productivity.

A review of the literature on productivity assessment shows that little is known about the effect of workplace literacy programs on job performance. This is due to the fact that many companies do not perform cost-benefit analyses on them. Of those who do, most do not publish the results. However, two measurements that can be used to evaluate job productivity are company records and job-related performance ratings. They can provide information on scrap rates, customer complaints, and supervisors' measurements of worker improvement.

For both educational and economic reasons, evaluation is becoming more critical for workplace literacy programs. As we move into the new information age, we will become even more reliant upon people with advanced knowledge and skills. Training is becoming an integral part of both educating and re-educating our nation. However, we must examine and evaluate our current educational practices and determine what works. Evaluation is needed to provide the most effective educational programs and also to justify their cost and existence.

We encourage you to make evaluation an integral part of your workplace literacy program, and help both practitioners and researchers learn more about how to create effective and successful workplace literacy programs.

References

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Bean, R., Partanen, J., Wright, F., & Aaronson, J. (1989). *Attrition in urban basic literacy programs and strategies to increase retention.* Pittsburgh, PA. (ERIC Document Reproduction Service No. ED 317 797)

Collino, G.E., Aderman, E.M., & Askov, E.N. (1988). *Literacy and Job Performance: A Perspective.* Pennsylvania State Institute for the Study of Adult Literacy.

Drew, R. A. & Mikulecky, L. J. (1988). How to Gather and Develop Job Specific Literacy Materials for Basic Skills Instruction. Bloomington, IN: The Office of Education and Training Resources, School of Education, Indiana University.

Greenville Technical College. (1993). *Training Opportunities Partnership Program.* "Final Performance Report, Evaluation Report, and Methodology Guide." Greenville Technical College, S.C. (ERIC Document Reproduction Service No. ED 354 361)

Haigler, K. (1990). *The Job Skills Education Program: An Experiment in Technology Transfer for Workplace Literacy.* A discussion paper prepared for the Work in America Institute, Harvard Club, New York, June 1990.

Hargroves, J. (1989). "The Basic Skills Crisis: One Bank Looks at Its Training Investment." *New England Economic Review.* September/October, pp. 58-68.

Kutner, M., Sherman, R., Webb, L. & Fisher, C. (1991*). A Review of the National Workplace Literacy Program.* Washington, DC: U.S. Department of Education, Office of Planning, Budget and Evaluation.

Lytle, S. L. (1990). *Living literacy: The practices and beliefs of adult learners.* Paper presented at the American Educational Research Association Annual Meeting, Boston, MA.

McBride, P. et al (1992). *Creating Custodial Classes: An Instructional Program Guide for Custodial Workers.* (ERIC Document Reproduction Service No. ED 353 378).

Mikulecky, L. J. (1985). *Literacy task analysis: Defining and measuring occupational literacy demands.* A paper presented at the annual meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED 262 206).

Mikulecky, L. & D'Adamo-Weinstein, L. (1991). "How Effective Are Workplace Literacy Programs?" In M.Taylor, G. Lewe, and J. Draper, (Eds.*) Basic Skills for the Workplace*. Toronto: Culture Concepts, Inc.

Mikulecky, L., & Lloyd, P. (1993). *Evaluating the impact of workplace literacy programs: Results and instruments from the NCAL workplace literacy impact project.* Philadelphia, PA: National Center on Adult Literacy, University of Pennsylvania. (ERIC Document Reproduction Service No. ED 348 579)

Mikulecky, L. & Lloyd, P. (1994). *Handbook of ideas for evaluating workplace literacy programs.* Bloomington, IN: Indiana University School of Education. (ERIC Document Reproduction Service No. ED 375 264).

Mikulecky, L. & Lloyd, P. (1995). *Evaluation of Workplace Literacy Programs: A Profile of Effective Instructional Practices.* Philadelphia, PA: National Center on Adult Literacy, University of Pennsylvania. (in press)

Mikulecky, L., Lloyd, P., Horwitz, L., Masker, S. & Siemantel, P. (1995). *A Review of Recent Workplace Literacy Programs and a Projection of Future Challenges.* Philadelphia, PA: National Center on Adult Literacy, University of Pennsylvania. (in press)

Mikulecky, L., Lloyd, P., Kirkley, J., & Oelker, J. (1995). *Developing and Evaluating Workplace Literacy Programs: A Handbook for Practitioners and Trainers.* Philadelphia, PA: National Center on Adult Literacy. (in press)

Mikulecky, L.J. & Strange, R. (1986). "Effective Literacy Training Programs for Adults in Business and Municipal Government." In J. Orasanu (Ed.) *Reading Comprehension: From Research to Practice.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Norback, J. Shaul. (in press). *Literacy Skills Analysis and Job Training.* Philadelphia, PA: National Center on Adult Literacy.

Norback, J. Shaul, Rosenfeld, M. & Wilson, S.T. (1994). *Job Literacy Analysis: A Practical Methodology for Use in Identifying Job-related Literacy Skills.* Princeton, NJ: Center for Skills Enhancement.

Philippi, J.W. (1988). "Matching Literacy to Job Training: Some Applications for Literacy Programs." *Journal of Reading*, 31 (7), 658-66.

Philippi, J.W. (1991). *Literacy at Work: The Workbook for Program Developers.* New York: Simon and Schuster.

Sheppeck, M.A. & Cohen, S.L. (1985). "Put a dollar value on your training programs." *Training and Development Journal*, 39, 59-62.

Stewart, T.A. (1994). "Your Company's Most Valuable Asset: Intellectual Capital." *Fortune*, Oct 3, 1994, *130*, (7), 68-74.

Sticht, T. G. (1982). *Basic Skills in Defense.* Alexandria, VA: Human Resources Research Organization.

U.S. Departments of Education and Labor (1988). *The Bottom Line: Basic Skills in the Workplace.* Washington, DC. Office of Public Information. Employment and Training Administration. Department of Labor.

Van Tilburg, E., & DuBois, J. E. (1989). *Literacy students' perceptions of successful participation in adult education: A cross-cultural approach through expectant values.* Paper presented at the Annual Adult Education Research Conference. Madison, WI. (ERIC Document Reproduction Service No. ED 321 030)